GraphAF: a Flow-based Autoregressive Model for Molecular Graph Generation

Year: 2020 | Citations: 489 | Authors: Chence Shi, Minkai Xu, Zhaocheng Zhu, Weinan Zhang, Ming Zhang

---

Abstract

Molecular graph generation is a fundamental problem for drug discovery and has been attracting growing attention. The problem is challenging since it requires not only generating chemically valid molecular structures but also optimizing their chemical properties in the meantime. Inspired by the recent progress in deep generative models, in this paper we propose a flow-based autoregressive model for graph generation called GraphAF. GraphAF combines the advantages of both autoregressive and flow-based approaches and enjoys: (1) high model flexibility for data density estimation; (2) efficient parallel computation for training; (3) an iterative sampling process, which allows leveraging chemical domain knowledge for valency checking. Experimental results show that GraphAF is able to generate 68% chemically valid molecules even without chemical knowledge rules and 100% valid molecules with chemical rules. The training process of GraphAF is two times faster than the existing state-of-the-art approach GCPN. After fine-tuning the model for goal-directed property optimization with reinforcement learning, GraphAF achieves state-of-the-art performance on both chemical property optimization and constrained property optimization.